

# MOReL : Model-Based Offline Reinforcement Learning

Year: 2020 | Citations: 742 | Authors: Rahul Kidambi, A. Rajeswaran, Praneeth Netrapalli, T. Joachims

---

## Abstract

In offline reinforcement learning (RL), the goal is to learn a highly rewarding policy based solely on a dataset of historical interactions with the environment. The ability to train RL policies offline can greatly expand the applicability of RL, its data efficiency, and its experimental velocity. Prior work in offline RL has been confined almost exclusively to model-free RL approaches. In this work, we present MOReL, an algorithmic framework for model-based offline RL. This framework consists of two steps: (a) learning a pessimistic MDP (P-MDP) using the offline dataset; and (b) learning a near-optimal policy in this P-MDP. The learned P-MDP has the property that for any policy, the performance in the real environment is approximately lower-bounded by the performance in the P-MDP. This enables it to serve as a good surrogate for purposes of policy evaluation and learning, and overcome common pitfalls of model-based RL like model exploitation. Theoretically, we show that MOReL is minimax optimal (up to log factors) for offline RL. Through experiments, we show that MOReL matches or exceeds state-of-the-art results in widely studied offline RL benchmarks. Moreover, the modular design of MOReL enables future advances in its components (e.g. generative modeling, uncertainty estimation, planning etc.) to directly translate into advances for offline RL.